

METHOD FOR LEARNING LINGUISTICALLY VALID WORD PRONUNCIATIONS FROM ACOUSTIC DATA

ABSTRACT OF THE DISCLOSURE

A computerized method is provided for generating pronunciations for words and storing the pronunciations in a pronunciation dictionary. The method includes graphing sets of initial pronunciations; thereafter in an ASR subsystem determining a highest-scoring set of initial pronunciations; generating sets of alternate pronunciations, wherein each set of alternate pronunciations includes the highest-scoring set of initial pronunciations with a lowest-probability phone of the highest-scoring initial pronunciation substituted with a unique-substitute phone; graphing the sets of alternate pronunciations; determining in the ASR subsystem a highest-scoring set of alternate pronunciations; and adding to a pronunciation dictionary the highest-scoring set of alternate pronunciations.

23310896 v1